1   1.   A computer program product for compressing data

2   files representative of an image document having color

3   information and/or graphical information, said software

4   product disposed on a computer readable medium comprising

5   instructions for causing a computer to:

6   provide a first image file at a first resolution and a

7   second image file at a second resolution of said document

8   with said second resolution being lower than said first

9   resolution;

10   process the first image file to convert the first image

11   file into a text file representation of the document;

12   compress said text file representation of the document

13   to provide a first compressed file;

14   process said second file to extract information from

15   the image representation of the document corresponding to

16   color information and graphics information;

17   compress the second file using a second compression

18   technique to provide a second compressed file containing

19   information corresponding to the image; and

20   store said first and second compressed files and color

21   information to provide a composite compressed file

22   corresponding to the document.

1   2.   The computer program product as recited in claim 1

2   wherein said instruction for causing the computer to provide

3   the image file comprises instructions for causing the

4   computer to scan the document at a first resolution to

5   provide said first image file at said first resolution and

6   scan the document at a second resolution to provide said

7   second image file at said second resolution.

1   3.   The method as recited in conjunction with claim 2

2   wherein said instruction to process the second file to

3   extract information corresponding to color information

4   comprises instructions for causing the computer to:

5   determine foreground colors corresponding to colors

6   associated with text portions of the document.

1   4.   The computer program product as recited in

2   conjunction with claim 3 further comprising instructions for

3   causing the computer to:

4   determine the foreground colors by causing the computer

5   to retrieve a plurality of samples of groups of pixels from

6   the low resolution image representation of the document and

7   from each one of said samples of pixels finding a pixel

8   corresponding to the minimum and maximum intensity of the

9   pixels in the sample;

10   and for each one of said samples

11   calculate a threshold value representative of the

12   document by averaging the minimum and maximum intensities

13   for each of the blocks;

14   determine a color associated with each one of the

15   blocks and the width of intensity of each one of the blocks

16   and

17   provide a data structure having an entry for each one

18   of said blocks corresponding to a foreground color and a

19   background color.

5. The computer program product as recited in conjunction with claim 1 further comprising instructions for causing the computer to detect and remove data in said image representation of the document which when viewed on a display depicts a crease over a portion of the document.

6. The computer program product as recited in conjunction with claim 5 comprises instructions for causing the computer to:

form a first mask to apply to the text representation of a document to mask out portions of said representation corresponding to image portions of the document; and

provide a second mask to apply to the low resolution image portion of the document to mask out portions of said document corresponding to text portions of the document.

7. A computer system including a computer software product for compressing data files representative of an image document, said document including color information and/or graphical information, said computer system including:

a processor to execute said software instructions;

a memory storing said software program;

a display which displays representations of said document;

said software product disposed on a computer readable medium comprising instructions for causing a computer to:

provide a first image file at a first resolution and a second image file at a second resolution of said

14   document with said second resolution being lower than said

15   first resolution;

16   process the first image file to convert the first

17   image file into a text file representation of the document;

18   compress said text file representation of the

19   document to provide a first compressed file;

20   process said second file to extract information

21   from the image representation of the document corresponding

22   to color information and graphics information;

23   compress the second file using a second

24   compression technique to provide a second compressed file

25   containing information corresponding to the image; and

26   store said first and second compressed files and

27   said color information to provide a composite compressed

28   file of the document.

              A

1   8.   The computer system as recited in claim 7 wherein

2   said instruction for causing the computer to provide the

3   first file comprises instructions for causing the computer

4   to scan the document at a first resolution to provide said

5   first image file at said first resolution and scan the

6   document at a second resolution to provide said second image

7   file at said second resolution.

1   9.   The computer system as recited in conjunction with

2   claim 8 wherein said instructions for causing the computer

3   to process the second file to extract information

4   corresponding to color information comprises instructions

5   for causing the computer to:

6   determine foreground color corresponding to colors

7   associated with text portions of the document.

1   10.   The computer system as recited in conjunction with
2   claim 9 further comprising instructions for causing the
3   computer to
4   determine the foreground colors by causing the computer
5   to retrieve a plurality of sample of pixels from the low
6   resolution image representation of a document and from each
7   one of said samples of pixels finding a pixel corresponding
8   to the minimum and maximum intensity of the pixels in the
9   sample;
10   and for each one of said samples
11   calculate a threshold value representative of the
12   document by averaging the minimum and maximum intensities
13   for each of the blocks;
14   determine a color associated with each one of the
15   blocks and the width of intensity of each one of the blocks;
16   and
17   provide a color data structure having an entry for each
18   one of said blocks corresponding to a foreground color and a
19   background color.

1   11.   The computer system as recited in conjunction with
2   claim 7 further comprising instructions for causing the
3   computer to:
4   detect and remove data in said image representation of
5   the document which when viewed on a display forms a crease
6   over a portion of the document.

1   12.   The computer system as recited in conjunction with

- 54 -

claim 11 comprises instructions for causing the computer to:

form a first mask to apply to the text representation of a document to mask out portions of said representation corresponding to image portions of the document; and

provide a second mask to apply to the low resolution image portion of the document to mask out portions of said document corresponding to text portions of the document.

15. A computer program product operable to determine a dominant background color associated with an image representation of a document comprises instructions for causing a computer to:

retrieve background color information associating a background color with each one of a plurality of samples of pixels representing the document;

filter said background colors to provide a target number of colors to represent the background colors;

apply a median cut analysis to the background color samples to filter said background samples into one of a plurality of boxes corresponding to said target number of colors;

sort said boxes by increasing volume;

sort a first portion of said boxes having the smallest amount of volume by decreasing intensity; and

determine the dominant background color as a color to represent the background of the document by the box having the lowest intensity.

16. A computer program product for decompressing a file containing image information and text information, said

3   program residing on a computer readable medium comprising

4   instructions for causing a computer to:

5   decompress the file containing image information and

6   text information into an image file and a text file;

7   allocate a target bit map to represent the decompressed

8   file;

9   insert the decompressed image information into the

10   target bit map at locations specified by information

11   contained in said file containing image information and text

12   information; and

13   insert text information into said target bit map in

14   accordance with positional information provided from the

15   decompressed text file.

1   17.   The computer program product of claim 16 further

2   comprising instructions for causing a computer to display

3   the reconstructed representation of the decompressed file on

4   a computer monitor.

1   18.   The computer program product of claim 16 further

2   comprising instructions for causing a computer to fill the

3   target output bit map with a color corresponding to a

4   dominant background color provided from color information in

5   the file.

1   19.   The computer program product as recited in

2   conjunction with claim 18 further comprising instructions

3   for causing a computer to apply a color to the text

4   information in the target bit map in accordance with said

5   color information provided from said file.

- 56 -

6   20. A computer system including a computer software

7   product, said computer system including:

8   a processor to execute said software instructions;

9   a memory storing said software program;

10   a display which displays representations of said

11   document, for compressing data files representative of an

12   image document, said document including color information

13   and/or graphical information;

14   said system operable to determine a dominant background

15   color associated with an image representation of a document

16   containing color and text information comprises instructions

17   for causing the computer to:

18   retrieve background color information associating a

19   background color with each one of a plurality of samples of

20   pixels representing the document;

21   filter said background colors to provide a target

22   number of colors to represent the background colors;

23   apply a median cut analysis on the background color

24   samples to assign said background samples into one of a

25   plurality of three dimensional boxes in R-B-G color space,

26   said plurality corresponding to said target number of

27   colors; and

28   determine the dominant background color as a color to

29   represent the background of the document by the three

30   dimensional box having the lowest intensity.

1   21. The computer system of claim 20 wherein the

2   computer program product further includes instructions for

3   causing the computer to

4   sort said three dimensional boxes by increasing volume;

5   and

6   sort a first portion of said three dimensional boxes

7   having the smallest amount of volume by decreasing intensity

8   to determine the dominate background color.

1   22. The computer of claim 21 wherein the computer

2   program product further includes instructions for causing

3   the computer to sort the three dimensional boxes in R-G-B

4   space by increasing distance (D) from the dominant

5   background color determined by:

6   D = sqrt((R1 - R2) * (R1 - R2) + (G1 - G2) * (G1 - G2)

7   + (B1 - B2) * (B1 - B2)).

1   23. The computer of claim 21 wherein the computer

2   program product further includes instructions for causing

3   the computer to estimate a width of the dominate backgroud

4   color by using a distance to one of the boxes in the first

5   portion of boxs as the width estimate.

1   24. The computer of claim 21 wherein the computer

2   program product further includes instructions for causing

3   the computer to map colors to the dominate background color.

1   25. A computer program product for a reading machine

2   to read words from a document allowed to a user and residing

3   on a computer readable medium comprising instructions for

4   causing a computer to:

5   decompress a file comprising image and text into an

6   image file and text file;

7   reconstruct the document from the decompressed image

- 58 -

8   file and text file;

9   apply color to the text in accordance with color

10 information provided from the file comprising image and text

11 and positional information provided from said text file;

12   display the reconstructed image representation of the

13 document on a computer monitor; and

14   manipulate the displayed reconstructed image

15 representation of the document by using positional

16 information in the text file.

1   26. The computer program product of claim 25 further

2   comprising instructions for causing the computer to scale

3   the positional information in the text file for manipulation

4   of the image in accordance with a resolution of the image

5   file and a resolution of the text file.

1   27. The computer program product as recited in

2   conjunction with claim 26 wherein said program further

3   comprises instructions for causing the computer to:

4   permit the user to select a document item from a

5   displayed reconstructed image representation of the

6   document; and

7   apply the converted text representation corresponding

8   to the document item to a speech synthesizer to cause the

9   computer to output a synthesized speech representation of

10 the document item.

1   28. The computer program product as recited in

2   conjunction with claim 27 further comprising computer

3   instructions for causing the coumputer to display image

- 59 -

4   representation of the document to be highlighted by applying

5   a highlighting indicia to the displayed reconstructed image

6   representation in accordance with scaled positional

7   information provided from the text file.

1   29. The computer program product of claim 28 further

2   comprising instructions for causing a computer to dissplay

3   an image representation of a document item to be highlighted

4   with a color by applying a color to the displayed image

5   representation in accordance with positional information

6   provided from the converted text file.

1   30. A method of compressing an image representation of

2   a document having color portions and text portions comprises

3   the steps of:

4   scanning a document to provide a first file at a first

5   resolution and a second file at a second, lower resolution;

6   converting the first file into a text file;

7   applying an auto-rotate filter to the first file to

8   correct said file for errors;

9   converting said high resolution image file into an

10   optical character recognition file having text information

11   and positional information corresponding to the text

12   information on the image document;

13   masking portions of said optical character recognition

14   file corresponding to portions of said document representing

15   graphical information associated with the document; and

16   compressing the unmasked portions of said optical

17   character recognition file to provide a compressed text

18   file;

19   applying a rotate filter to the second file to correct

20 errors in said second file;

21   determining from said second file foreground colors

22 associated with each of the sections of said document and

23 background colors associated with each portion of said

24 document;

25   determining from said background colors a dominant

26 background color that best represents the background color

27 of the document;

28   masking portions of said document not corresponding to

29 the graphical portions of the document; and

30   compressing said unmasked portions to provide a second

31 file corresponding to graphical portions of the document and

32 storing said color information, and said first and second

33 files as a composite file.

29
3Ɩ.   A reading machine comprising:

    a computer, said computer comprising:

        a processor;

        a computer monitor for displaying the image

5 representation of a document having text and graphical image

6 content;

7   a mass storage device, said storage device including

8 software comprising instructions for causing the computer

9 to:

10   decompress a file into an image file and text

11   file;

12   reconstruct the document from the

13   decompressed image file and text file and apply

14   color to the text in accordance with color

15　　　　　information and positional information provided

16　　　　　from said text file; and

17　　　　　　　display the reconstructed image

18　　　　　representation of the document on a computer

19　　　　　monitor and manipulate the displayed reconstructed

20　　　　　image representation of the document by using

21　　　　　positional information in the text file.

30

1　　　32. A reading machine system of claim 31 further

2　comprising instructions for causing the computer to scale

3　the positional information in the text file for manipulation

4　of the image in accordance with a resolution of the image

5　file and a resolution of the text file.

31

1　　　33. A reading machine system as recited in conjunction

2　with claim 32 wherein said program further comprises

3　instructions for causing the computer to:

4　　　permit the user to select a document item from a

5　displayed reconstructed image representation of the

6　document; and

7　　　apply the converted text representation corresponding

8　to the document item to a speech synthesizer to cause the

9　computer to output a synthesized speech representation of

10　the document item.

32

1　　　34. A reading machine system as recited in conjunction

2　with claim 33 further comprising instructions for causing

3　the computer to display image representation of the document

4　to be highlighted by applying a highlighting indicia to the

5　displayed reconstructed image representation in accordance

6   with scaled positional information provided from the text

7   file.

35.   The reading machine system of claim 34 further

comprising instructions for causing a computer to display an

image representation of a document item to be highlighted

with a color by applying a color to the displayed image

representation in accordance with positional information

provided from the converted text file.

36.   A reading system comprising:

a computer, said computer comprising:

        a processor;

        a computer monitor for displaying an image

representation of a document having text and graphic or

image content information;

        a mass storage device, said storage device

including a computer program product for decompressing

a file containing image information and text

information, said program residing on a computer

readable medium comprising instructions for causing a

computer to:

            decompress a file into an image file and a

        text file;

            allocate a target bit map to represent the

        decompressed file;

            insert the decompressed image information

        into the target bit map at locations specified by

        information contained in said file; and

            insert text information into said target bit

21 map in accordance with positional information

22 provided from the decompressed text file.

1 37. The reading system of claim 36 further comprising

2 instructions for causing a computer to display the

3 reconstructed representation of the decompressed file on a

4 computer monitor.

1 38. The reading system of claim 36 further comprising

2 instructions for causing a computer to fill the target

3 output bit map with a color corresponding to a dominant

4 background color provided from color information in the

5 file.

1 39. The reading system as recited in conjunction with

2 claim 38 further comprising instructions for causing a

3 computer to apply a color to the text information in the

4 target bit map in accordance with said color information

5 provided from said file.

1 40. A computer program product for detecting a crease

2 in an input color image of a document, said software product

3 disposed on a computer readable medium comprising

4 instructions for causing a computer to:

5 retrieved data corresponding to portions of the page

6 that divides the page into sections;

7 and for each one of the sections

8 retrieve a one-dimensional array of pixels with a

9 length equal to the width of the portion; and

10   determine for each of the one-dimensional array of
11   pixels whether the array has an intensity minima, by
12   measuring a distance from the center of an assumed intensity
13   minima out to a dominant background color.

1   41. The computer program product of claim 40 wherein
2   for any slices that contain no dominant background color
3   pixels in the appropriate direction such samples are
4   considered to be invalid and are discarded.

1   42. The computer program product of claim 41 wherein
2   for those samples that have a dominant background color at
3   the appropriate location, a center for the crease is
4   determined by averaging the centers of the best slices.

1   43. The computer program product of claim 42 wherein
2   the first average of the centers of all the valid slices are
3   sorted by increasing distance from the first average and the
4   average is recomputed using only the centers of the highest
5   (NSLICES/2)+1, where (NSLICES) is the number of slices.

6   44. The computer program product of claim 43 wherein
7   the whole area is considered to be invalid if there are less
8   than (NSLICES/2)+1 valid slices.

1   45. The computer program product of claim 44 wherein a
2   composite width is assigned for the area crease as the
3   minimum area slice width, and a composite vector of
4   intensities for each slice is constructed from the center
5   point of the crease to the near dominant background color

6   point for the slice.

1   46.   The computer program product of claim 45 wherein

2   an array corresponding to the composite vector of

3   intensities for each slice is filled in as follows:

4   for a "center" area;

5   define slice(I) to be the pixel in a slice that is I

6   number of pixels from the center in the direction of the

7   near-DBC point; and

8   fill in the array;

9   array[i] = average of intensities of the slice[i]

10   pixels for the valid slices; and

11   iterate over I from the center out to the near-DBC

12   point as:

13   array[i] = maximum of array[i] and array[i - 1]

14   for each side of the crease, producing two arrays.

1   47.   The computer program product of claim 45 further

2   comprising the step of;

3   assigning a quality to each area of the page with the

4   quality being equal to the width of the crease found or an

5   invalid crease indicator if the area/crease fails to qualify

6   as a crease

7   if there are less than   (NSLICES/2)+1 valid slices, or

8   the width is below a minimum crease width, or if the

9   majority of centerpoints used to construct the average

10   centerpoint are not within a constant horizontal distance or

11   one another or if the vector of intensities appears concave.

1   48.   The computer program product of claim 47 wherein

2   the crease with the highest quality is determined as the

3   crease for the page.

1   49. A computer program product for removing a crease

2   stored on a computer readable media, comprises instructions

3   for causing a computer to:

4   set all pixels to the outside of the center portion of

5   a left or right side of the image crease to a dominant

6   background color.

1   50. The computer program product of claim 49 further

2   comprising instructions for causing a computer to:

3   set all pixels within a fixed distance to the dominant

4   background color unless the width of the crease is less than

5   a predetermined value; and

6   bleach all other pixels within the width of the crease

7   between the center point and the near dominant background

8   color point.

1   51. The computer program of claim 50 wherein the

2   instruction for causing the computer to bleach comprises

3   instructions for causing the computer to:

4   define array[i] to be the intensity in the creases's

5   intensity vector at a distance I pixels from the center;

6   define image[y][x] to be the pixel in the image x

7   pixels horizontally and y pixels vertically from the upper-

8   left corner;

9   define center to be the center of the crease and width

10   to be its width;

11   define intensity(pixel) to be a function that returns

the intensity of a pixel;

        for a left-side crease, iterate over y, for each row in the image, iterate over I from a fixed distance over crease width:

    if ((intensity(image[y][center + I] + (intensity(DBC) - array[i]))) > (0.90 * intensity(DBC)))

        set image[y][center + I] to DBC and set the corresponding pixels in the B/W image to white.

52. A computer program product for detecting and removing a crease in an input color image of a document, said product disposed on a computer readable medium comprising instructions for causing a computer to:

        retrieved data corresponding to portions of the page that divides the page into sections; and for each section

        retrieve a one-dimensional array of pixels with a length equal to the width of the portion;

        determine for each of the one-dimensional array of pixels whether the array has an intensity minima, by measuring a distance from the center of an assumed intensity minima out to a dominant background color;

        set all pixels to the outside of the center portion of a left or right side of the image crease to a dominant background color;

        set all pixels within a fixed distance to the dominant background color unless the width of the crease is less than a predetermined value; and

        bleach all other pixels within the width of the crease between the center point and the near dominant background color point.

add BE

- 68 -